Peer reviewed version

Link to published version (if available):
10.1177/1475921716651809

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Prediction of Landing Gear Loads Using Machine Learning Techniques

Geoffrey Holmes[1], Pia Sartor[2], Stephen Reed[3], Paul Southern[4], Keith Worden[1] and Elizabeth Cross[1]

[1]Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield,

Mappin Street, Sheffield, S1 3JD, UK

[2]Department of Aerospace Engineering, University of Bristol, University Walk, Bristol, BS8 1TR, UK

[3] dstl, Porton Down, Salisbury, SP4 0JQ, UK

[4]Messier-Bugatti-Dowty, Messier-Bugatti-Dowty, Cheltenham Rd East, Gloucester, GL2 9QH

+44 (0) 114 222 7846 | e.j.cross@sheffield.ac.uk

Keywords: Landing Gear, Loads, Machine Learning, Gaussian Process Regression

ABSTRACT

This paper investigates the feasibility of using machine learning algorithms to predict the loads experienced by a landing gear during landing. For this purpose, results on drop test data and flight test data will be examined. This paper will focus on the use of Gaussian Process regression for the prediction of loads on components of a landing gear. For the learning task, comprehensive measurement data from drop tests are available. These include measurements of strains at key locations, such as the on the side-stay and torque link, as well as acceleration measurements of the drop carriage and the gear itself, measurements of shock absorber travel, tyre closure, shock absorber pressure and wheel speed. Ground-to-tyre loads are also available through measurements made with a drop test ground reaction platform. The aim is to train the GP to predict load at a particular location from other available measurements, such as accelerations, or measurements of the shock absorber. If models can be successfully trained, then future load patterns may be predicted using only these measurements. The ultimate aim is to produce an accurate model that can predict the load at a number of locations across the landing gear by using measurements that are readily available, or may be measured more

easily than directly measuring strain on the gear itself (for example, these may be measurements already available on the aircraft, or from a small number of sensors attached to the gear). The drop test data models provide a positive feasibility test which is the basis for moving on to the critical task of prediction on flight test data. For this, a wide range of available flight test measurements is considered for potential model inputs (excluding strain measurements themselves), before attempting to refine the model or use a smaller number of measurements for the prediction.

I. INTRODUCTION

The landing gear of an aircraft is a unique component because it is both a structure and a system. It is a complex system with controlled articulation, multiple axes of energy absorption, and it is the sole structure supporting the aircraft when on the ground. It must bear extreme and varying loads when an aircraft manoeuvres on the ground during taxi, take-off and landing, yet it must be lightweight and compact because it is stowed and unused during the majority of an aircraft's flight [1].

The significant difference between the aircraft structure and landing gear is reflected in the difference between their design and approval methodologies. Many airframe designs use a "damage tolerant" design philosophy, which assumes that cracks exist in the structure; the structure is designed to retain adequate strength until the crack is detected and corrective action is taken. However, landing gear use a "safe life" design philosophy in which the component is designed to a specific service life and is removed from service before this elapsed time so that the probability of failure is remote [2]. This approach is followed because in most applications, landing gear have no structural redundancy. Furthermore, landing gear are predominantly manufactured from very high strength (but relatively low toughness) steel and titanium alloys as they must be able to withstand high loads but have minimum weight and size [3]. In high strength materials, crack growth is very rapid and critical crack sizes may not be detectable. Therefore, landing gear are designed to an assumed loads usage spectrum. However, how closely the actual landing gear usage matches the assumed loads spectrum is unknown.

Taking this into account, there are several areas of interest in aircraft landing gear structural health monitoring: understanding of the current operating environment of landing gear in order to allow for an

improvement in the assumed loads usage spectrum used in the evaluation of fatigue design criteria, surveillance of the landing gear fleet in order to detect overload occurrences (and equally to indicate which occurrences were not overloads), and ultimately to allow the certification of the landing gear to be based on the actual life experienced in service [1].

Different approaches have be taken to determine in-service landing gear loads, such as the use of kinematics (accelerations, velocities and displacements) or the use of other measurements such as pressure or strain. Generally work has focused on developing inertial measurement systems to provide kinematic flight parameter data that can be used to calculate landing gear loads using physics-based models [4-6]. Other work focused on directly measuring the forces on the landing gear structure has involved directly instrumenting the landing gear with strain gauges [7] or by placing a transducer, such as a load pin, in a landing gear load path [8, 9]. The development of aircraft weight and balance systems, which employ strain gauge based or fibre Bragg grating transducers in the landing gear axle, also provided a means for force measurements on the landing gear [10-12]. Additionally, landing gear shock absorber gas and hydraulic fluid measurements have been used for determining the loads in a single axis of the landing gear [13, 14]. In all of these methods, however, the additional weight and reliability issues of these systems have limited their use.

If landing gear loads can be inferred from flight parameters through machine learning techniques, it will avoid the undesirable scenario of having to place additional systems on the aircraft to determine these loads. Researchers in the aerospace industry have used machine learning techniques such as artificial neural networks (ANNs) to determine the loads on other parts of the aircraft structures. Significant work in this area is by Azzam et al. [15, 16] and Wallace et al. [17], who used a mathematical network to predict loads at a number of structural locations for the Tornado combat aircraft, over a large number of test flights. Azzam et al. [16] also described the development of methods to predict damage in helicopter components and the prediction of high-frequency events, such as buffet loading on the fin of a fixed-wing aircraft. The work reported in the current paper was inspired by the work of Reed [18, 19], who developed an ANN-based parametric fatigue monitor for the wing and tailplane of a military trainer aircraft and the wing of a combat aircraft.

The machine learning approach in this paper utilises Gaussian Process (GP) regression along with greedy algorithms for optimisation in order to predict loads on components of the landing gear structure. The use of GPs is a growing area of interest in many disciplines where they are employed as a sophisticated nonparametric Bayesian approach to regression and classification problems. Most recently in the field of SHM, GP regression has been used for prediction of crack growth in aluminium specimens [20] and as a predictor for key features on a suspension bridge [20]. An example of the use of GPs for acoustic source location [22] is applied in the context of landing gear SHM in [23].

One of the key benefits of using machine learning techniques is the potential for minimising the additional aircraft instrumentation required. Ideally, all of the required information will be available from the aircraft systems. An additional benefit of this approach is that it could easily be expanded to the aircraft maintenance monitoring system. This leads to a method of in-service loads monitoring that is advantageous in terms of weight, system complexity, reliability and cost.

Preliminary case studies have shown the feasibility of predicting landing gear loads using drop test data and flight test data. In [24], Multi-Layer Perceptron (MLP) and Bayesian MLP neural networks were developed using drop test data to predict landing gear side-stay loads. In [25], GPs were developed using flight test data to predict vertical ground-to-tyre loads.

This paper extends the previous work of the authors to use machine learning technique to predict landing gear ground-to-tyre loads and internal loads from drop test data and flight test data. In Section II, the main landing gear structure and typical landing loads are first described. The mathematical modelling approach that is used in the development of the GP regression models to predict the landing gear loads is then discussed in Section III. This study also includes the optimisation of the input parameter selection using a greedy algorithm. The results of applying such models to the drop test and flight test data are then presented in Section IV and finally, future plans for model development and testing are discussed in Section V.

II. LANDING GEAR LOADS

Figure 1 shows a typical telescopic port main landing gear (MLG) structure as well as a drop test rig. Drop tests are performed as part of the aircraft certification to verify the dynamic compression damping and energy
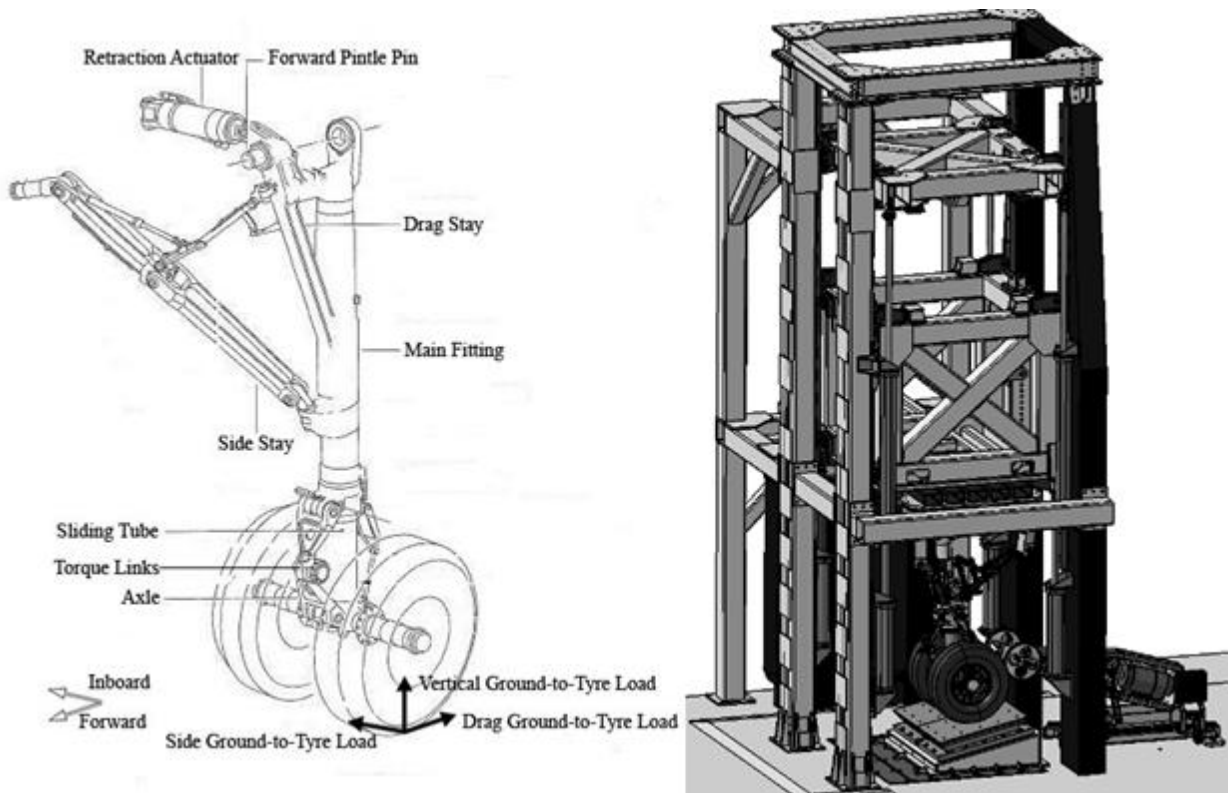
Figure 1. Main Landing Gear Structure and Drop Test Rig [27]

absorption characteristics of the landing gear shock absorber on landing [26]. In a drop test, the landing gear is mounted in a fixture that geometrically represents the aircraft landing gear attachment structure. The landing gear is dropped from various heights onto a ground reaction platform. The drop height is set to achieve the required vertical descent velocity. The correct proportion of landing weight is supported by the moving carriage and wing lift is simulated by upward acting jacks. Prior to the drop, the wheels are spun-up in the reverse rotation to simulate the aircraft forward landing speed. Landing attitude is varied by angling the landing gear in the fixture or angling the ground reaction platform. The surface of the ground reaction platform simulates the tyre-runway friction. The ground-to-tyre loads are measured using loads cells in the ground reaction platform. The landing gear side-stay and torque links are fitted with strain gauges and calibrated in order to directly measure loads.

Figure 2 illustrates the landing dynamics of the MLG in a landing. On contact with the ground reaction platform, the drag force deforms the landing gear aft and stores energy in the structure. When the tyre

velocity reaches zero, the strain energy stored in the rearward deformation produces a spring-back. The landing gear oscillates until the structural damping reduces the stored energy to zero [28]. Also during this time, there is an increasing vertical ground-to-tyre load, which is a function of the gas spring, oil damping (related to the square of the vertical descent velocity) and bearing friction. The shock absorber continues to close until all the vertical energy has been absorbed and then it partially recoils [29]. The landing gear side stay acts in the drag and side plane to prevent sideways movement of the landing gear and also forms an attachment to the aircraft structure. The torque links maintain the alignment of the axle.
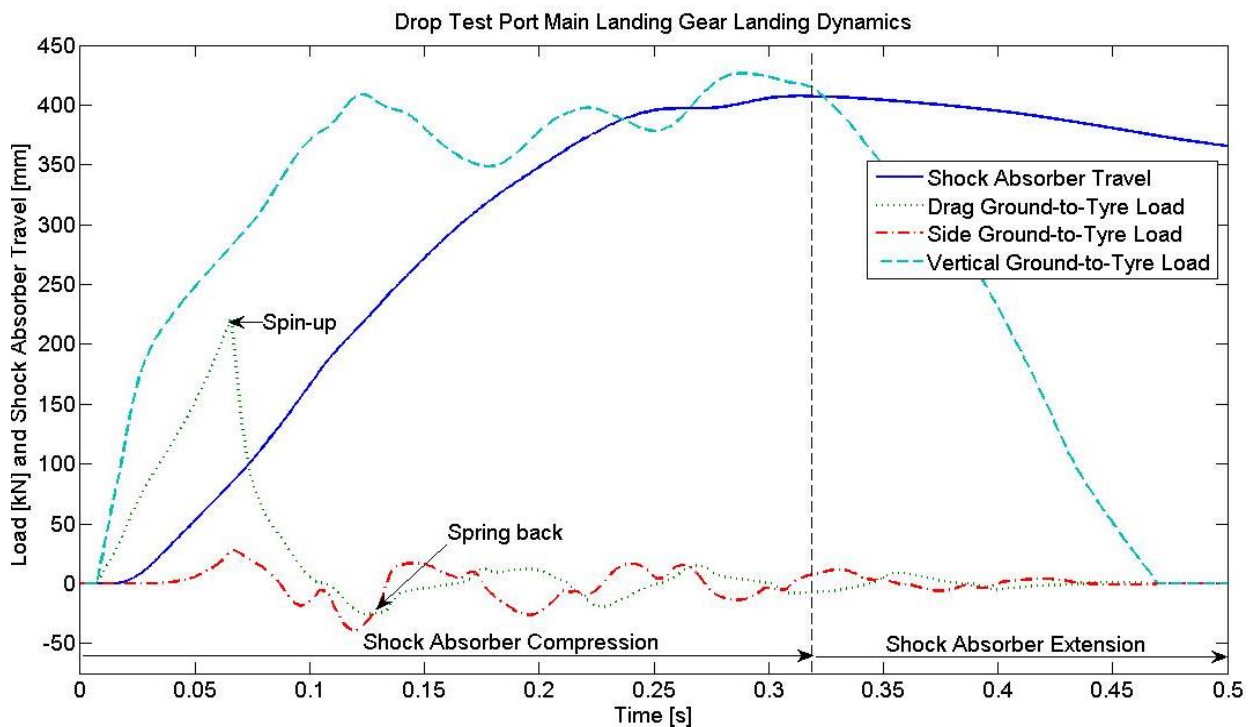
Figure 2. Example of Main Landing Gear Landing Dynamics

In this paper, the drop test data were collected from 21 individual drop tests and cover four different test set-ups in which the vertical descent velocity, wheel speed and ground friction were varied. Flight test data from a commercial aircraft undergoing typical landings were also obtained for this study. Examples of the types of measurements available as model inputs for the drop test and flight test data are listed in Table 1 along with the loads from the strain measurements used as model targets (i.e. the loads to be predicted) listed in Table 2.

TABLE I. DROP TEST AND FLIGHT TEST MODEL INPUTS

Drop Test	Flight Test
Carriage vertical descent velocity	Aircraft velocity (longitudinal, lateral, vertical)
Carriage acceleration	Aircraft accelerations at cabin centre of gravity
Carriage travel	Aircraft altitude
Carriage orientation is fixed	Aircraft pitch, roll, yaw
MLG main fitting acceleration	MLG main fitting accelerations
MLG sliding tube acceleration	-
MLG axle acceleration	MLG axle accelerations
Shock absorber gas pressure	Shock absorber gas pressures
Shock absorber travel	Shock absorber travel
Wheel speeds	Wheel speeds
Tyre closure	-
-	Brake Pressures

**TABLE II. DROP TEST AND FLIGHT TEST MEASURED OUTPUTS
WHICH ARE CANDIDATES FOR MODELLING**

Drop Test	Flight Test
Side-Stay Load	-
Torque Link Load	-
Drag Ground-to-Tyre Load	-
Vertical Ground-to-Tyre Load	MLG Vertical Load
Side Ground-to-Tyre Load	-

TABLE III. DETAILS OF DROP TEST DATA CONDITIONS

Condition	No. of tests	Vertical impact speed (m/s)	Ground friction
1	6	1.22	Normal
2	6	1.83	Normal
3	6	2.44	Normal
4	3	1.9	Increased

III.   THEORY

A. Gaussian Process Regression

GP regression, unlike classical maximum likelihood approaches based on a parameterised model form,

considers a family of functions that fit to a training data set and provides a predictive distribution as opposed

to a single crisp prediction for a given input. From this predictive distribution, a mean prediction and associated confidence intervals can be obtained. This nonparametric approach for regression has the benefit that model complexity is not limited by a set functional form. An additional benefit of using GPs lies in their compactness; the computations necessary for GP regression are simplified by the fact that a distribution directly over candidate functions can be defined, rather than over the parameters of a predefined function (as would be necessary for a Bayesian neural network for example). The full details of the computations involved in GP regression can be found in [30]. Very brief details will be given here in order to describe the user choices made.

To define and train a GP, a mean function, $m(x)$, and covariance function, $k(x, x')$, must first be specified. The mean function defines the expected prediction at any input location in the absence of any training data. Commonly, because little is known about the data at this stage, and for simplification purposes, the mean function is set to zero. This practise was followed in this study. The covariance function determines the covariance of the two predictions at any two specified points. A common choice of covariance function is the squared exponential, which has a general form [30];

$$k(x_p, x_q) = \sigma_f^2 exp\left(\frac{-1}{2l^2}|x_p - x_q|^2\right) + \sigma_n^2 \delta_{pq} \tag{1}$$

Where, x_p, x_q are input values, σ_f^2, σ_n^2 are signal and noise variances respectively and l is a characteristic length scale. Equation (1) specifies that the covariance function will vary smoothly from point to point and that function values at similar inputs will be highly correlated. The rate at which this correlation decays with distance is controlled by the choice of the length scale, l. The key step in training a GP is the specification of the hyper-parameters σ_f, σ_n, l. Within the machine learning community the most commonly used approach for optimising these hyper-parameters is to use a maximum likelihood approach to maximise the marginal likelihood of the training data [30]. That approach was followed here.

B. Model Performance

In this work, in order to quantify model performance, the mean squared error (MSE) is utilised as a measure of model fitness. Specifically, a normalised mean squared error will be used, which is defined here as:

$$MSE = \frac{100 \sum (model\ errors)^2}{n\ (\sigma[data])^2} \qquad (2)$$

Where n is the number of data points at which prediction were made and σ denotes standard deviation. This MSE is frequently used in structural system identification and has the property that, if the mean of the data is used as the model prediction, the MSE will be 100%. With this normalisation, values of MSE below 100% are indicative of captured correlation.

C. Optimisation of Input Parameter Selection

In the results section following, results are presented which made use of all available measurements as input variables for the GP. However, subset selection for model inputs was also investigated for the possibility that model fidelity would improve. There are a number of reasons why a smaller set of variables may yield better results than the full set. Firstly, there may be some variables that are uncorrelated with the output of interest; such variables simply increase the dimension of the problem and such 'uninformative' input variables should clearly be removed. Secondly, a similar issue also arises if input variables are correlated with the output, but also highly correlated with each other; in this case one does not wish to remove all the set, only enough to eliminate linear dependence between variables. This is a particularly important issue when a large number of possible input variables are available for inclusion in the model, as is the case for the flight test data in this study. Finally, adding input variables can decrease model fidelity if the basic variables are correlated with the output, but carrying so much noise that it dominates the predictions when passed through the model. The latter issue is partially addressed here by the filtering operations discussed below.

In order to select a useful subset of inputs for the GP training a greedy algorithm was used. Greedy algorithms are known to be sub-optimal but in many situations they provide a good first attempt at optimisation. They are particularly useful when model estimation is computationally intensive, as is the case for GP models using large amounts of training data, because only a limited number of models are tested. An example of the use of a greedy algorithm for subset selection in an SHM context can be found in [31]. In the present study an additive version of greedy algorithm was used: One begins with an empty set of input variables. One then fits N models with a single input. If the ith model has lowest MSE, input i is selected for inclusion in the variable

subset. At the next iteration, one fits $N - 1$ models with two input variables, the one already selected and one of the remaining unchosen inputs, and adds to the subset the combination which gives lowest MSE, and so on. The algorithm is terminated when the error begins to rise when new input variables are added.

IV. RESULTS

The following section describes the results from predicting the landing gear loads using drop test data and flight test data. Following established best practice in machine learning, the trained GPs are tested on data not included in the training set, and the results shown here are predictions on the test set.

A. Drop Test Data

For this study, drop test data were available from four different testing conditions (with differing drop heights, wheel spin up conditions and ground friction). For each test condition, data were collected from multiple drops. As discussed previously, the landing gear used was instrumented with strain gauges and accelerometers at key locations, and measurements of other variables such as shock absorber pressure were also available.

Initial studies focused on the prediction of the load measured on the side-stay, with all other measured parameters (excluding other strain measurements) considered as possible model inputs. In order to create a model that is able to predict load across different test conditions, data from two drops in every test condition were collated to create a training data set. One limitation of the GP implementation used here is the fact that it is necessary to invert an $N \times N$ matrix, where N is the number of points in of the training data set, although ways to circumnavigate this are now emerging [33]. In order to avoid very large matrix inversions, the size of the training data used to condition the GP was generally limited to be under 5000 samples, and to do this, subsampling a quarter of the available training data set was necessary. This was done by selecting the every 4th data point from the data in the time-series order in which it was collected. Each data point is a vector containing the input measurements at that point in time and the selected data points were passed to Equation (1) to form the elements of the covariance matrix [30].

The models were trained on data from two drops in every test condition using all available input parameters (see Table I, left column) and were able to predict the general trend of side-stay load in unseen data to a reasonable degree of accuracy as reported below. However, as will be shown later in this paper, the prediction accuracy can be much improved by considering filtering and subset selection of the input parameters. A typical prediction of side-stay load on an unseen data set is shown in Figure 3. From this figure one can see that the general trend of the load can be predicted successfully. It should be noted, however, that the prediction has a higher noise content than the target and that at times the measured load is outside the confidence interval for the prediction.
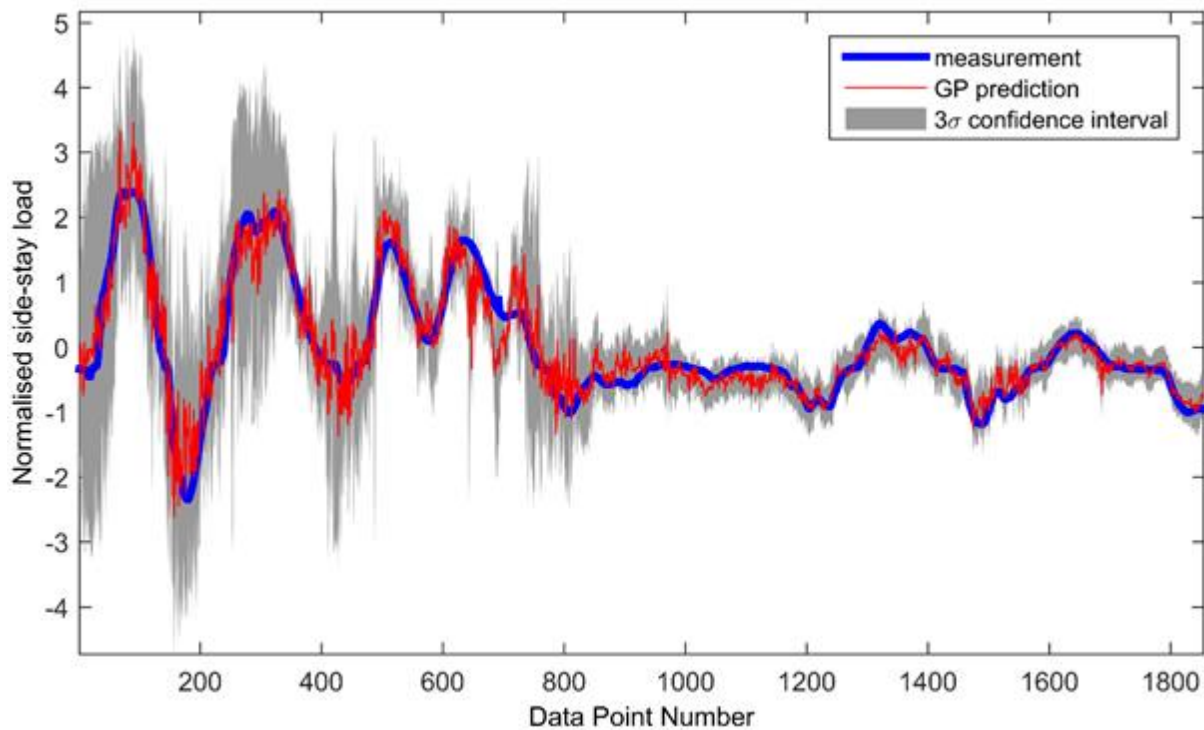
Figure 3. Prediction of side-stay load on an unseen drop test

The MSE for the drop load prediction shown in Figure 3 is 13.1%. In this and the following figures, the data point number corresponds to the relative position of the datum in the time series in which the data was collected. The range of MSE values across the four conditions was 13.2%-19.9% with a mean of 16.2%. Looking at the figure, it is likely that the noise content of the prediction contributes significantly to the MSE.

In order to improve model prediction fidelity, filtering of the input data set was considered. This can be justified by the observation that, in the dataset the noise content of any acceleration measurements is much higher than that of the load from the strain measurements. Figure 4 compares the power spectral density of a measured acceleration on the main fitting of the landing gear and load time history for the test data set shown in Figure 3. In this figure one can see that while the spectra match up to approximately 50 Hz, the accelerations have considerably more high frequency content than the load from the strain measurements.
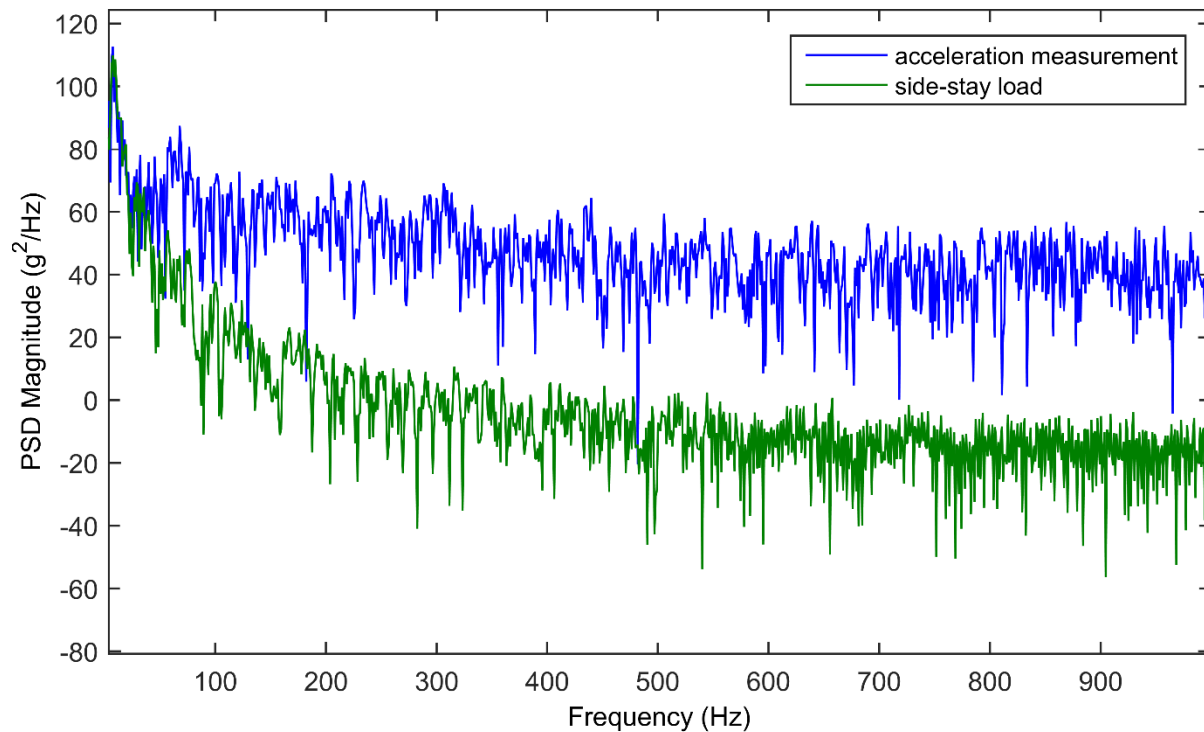

Figure 4. Comparison of power spectral density for a typical acceleration measurement on landing gear and side-stay load (normalised signals).

In order to reduce the discrepancy between acceleration and load, a low-pass filter with a cut-off of 225 Hz was applied to all acceleration measurements. This value was chosen in anticipation of fitting dynamic models, to allow up to third order interactions between frequencies. Rather than filter data from individual drops, concatenated drop data were filtered. This is because, in order to achieve a zero-phase filter, it was necessary to run the data through the filters both backwards and forwards. This operation engenders a filter transient at both the start and end of the filtered record. It was observed that the transients at the boundary points

in the concatenated data were less severe than when individual drops were filtered; this is presumably because each drop returns close to equilibrium by the end of each test record.

When applying a low-pass filter to the model inputs before training and testing, it was found that the GP predictions had lower MSE values. Figure 5 shows the GP predictions of the side-stay load when using filtered inputs for the same test data as shown in Figure 3. The MSE is a much improved 7.1%. The average MSE across all test data when using filtered inputs is 7.2% (this compares to an MSE of 16.2% when using unfiltered inputs as reported above).
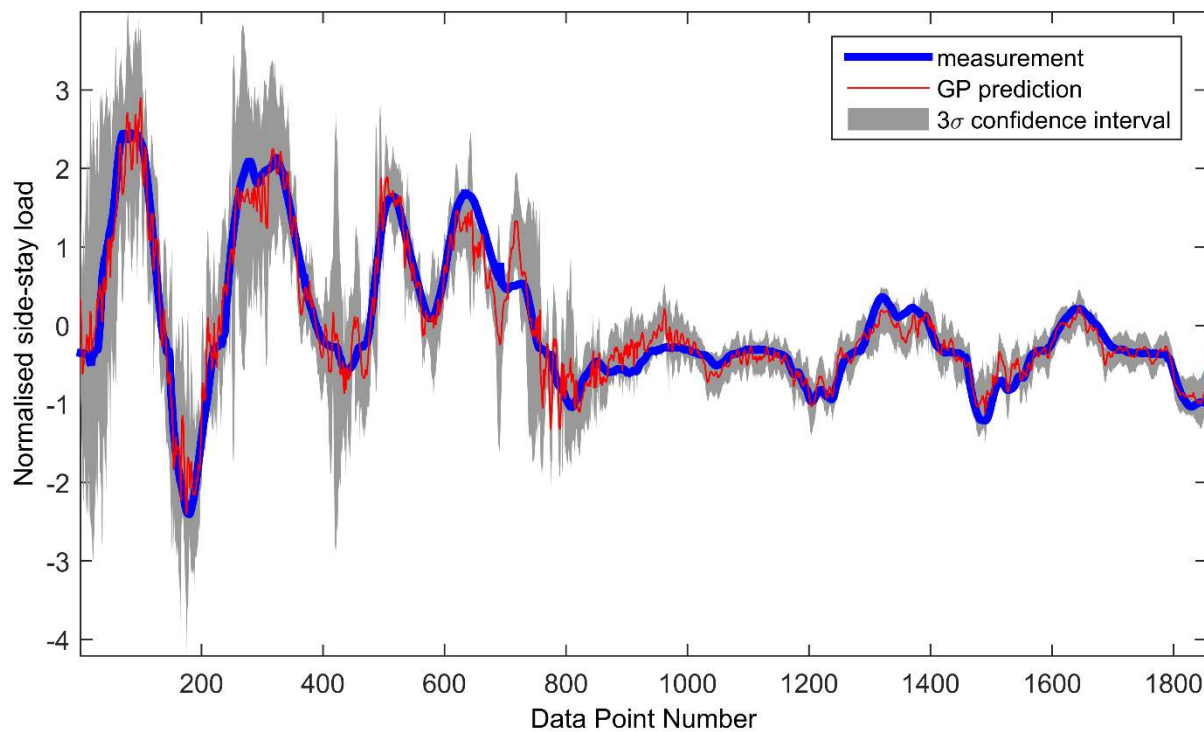
Figure 5. GP prediction of side-stay load using filtered inputs

The greedy algorithm with the insert strategy was run on the drop test data example considered in the last section. The stopping criterion for the algorithm was to stop if no decrease in MSE could be gained by adding an additional variable. The target variable was again chosen as the side-stay load and the candidate pool of input variables was made up of all inputs used in the models discussed in the last section (i.e. the filtered versions of those listed for the Drop Test data in Table I).

Table IV compares the average test MSE across different drop test conditions when using a GP with a full input set against one using inputs selected by the greedy algorithm. Studying Table IV one can see that in two out of four test conditions, the GP with inputs selected by the greedy algorithm is more successful than the

Table IV. Comparison between model errors of GPs trained on complete input parameter set and the input parameter sets selected by the greedy algorithm (filtered inputs). For each load predicted the lowest MSE is highlighted

Side-Stay Load Prediction – Test MSE (%)		
Drop Test Condition	**All Inputs**	**Greedy Selection**
1	6.6	10.1
2	6.8	7.1
3	7.8	5.2
4	7.7	6.0
Mean	7.2	7.1

GP with a full input set. The predictions in drop conditions 1 and 2 are better, however, if a full input set is used. This illustrates the fallibility of a greedy algorithm – the stopping criterion has in this case prevented potentially useful combinations of input parameters from being used.
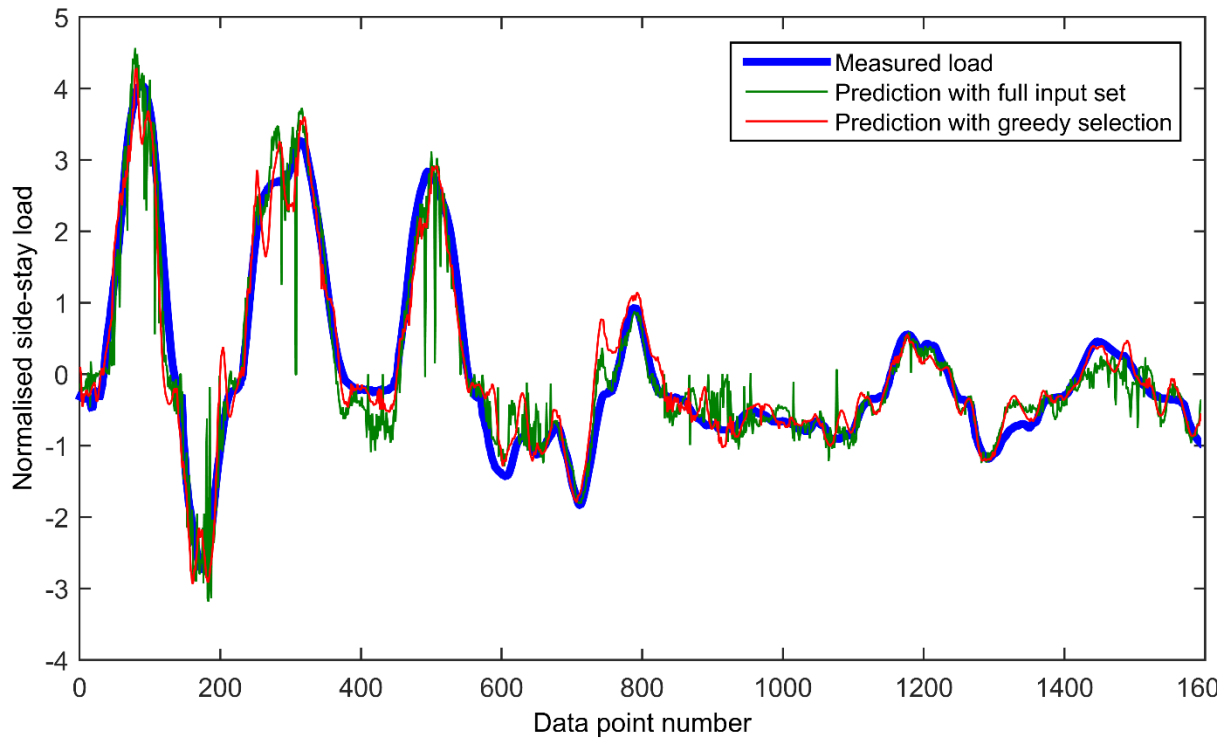
Figure 6. Comparison of model predictions using full input parameter set and greedy selection. Confidence intervals have been omitted here for clarity.

Figure 6 shows a comparison of model predictions in drop condition 3 for a GP with a full input set and one with inputs selected by the greedy algorithm. One can see that the reduced model error when using the greedy input selection could again be attributed to having a lower noise content than the prediction using all inputs.

Although only small gains have been seen here when using a greedy algorithm, such optimisation routines become much more valuable where the number of possible input variables is much larger – this will be illustrated in later sections where flight test data are discussed.

Table V. MSE of GP predictions across all drop test data conditions
model inputs selected by greedy algorithm

| Load Predicted	Test MSE (%)	
	Mean	Range
Side-stay load	7.1	5.2-10.1
Torque-link load	15.2	10.3-20.7
Drag ground-to-tyre load (ground reaction platform)	4.7	2.8-7.0
Side ground-to-tyre load (ground reaction platform)	14.3	6.8-22.3
Vertical ground-to-tyre load (ground reaction platform)	0.61	0.22-1.33

In the previous sections, accurate predictions of side-stay load on unseen data sets have been achieved with GP regression, filtering of input variables and a set of input variables selected by a greedy algorithm. Taking the same approach, trained GPs are also able to predict load at the torque link and ground-to-tyre loads to a high degree of accuracy. Table V shows the average MSE error across all test data (including different conditions) for a number of different measured loads. Typical predictions of each load are shown in Figures 7 to 10.

From Table V and Figures 7 to 10 one can see that the GPs are able to predict landing gear torque link and ground-to-tyre loads to a good degree of accuracy. The best predictions are achieved when considering vertical ground-to-tyre load, here the predictions are very accurate with a low MSE. Torque-link load and side ground to tyre load are more difficult to predict and show higher MSE values, although Figures 7 and 9 show that reasonable predictions are made on unseen (test) data.
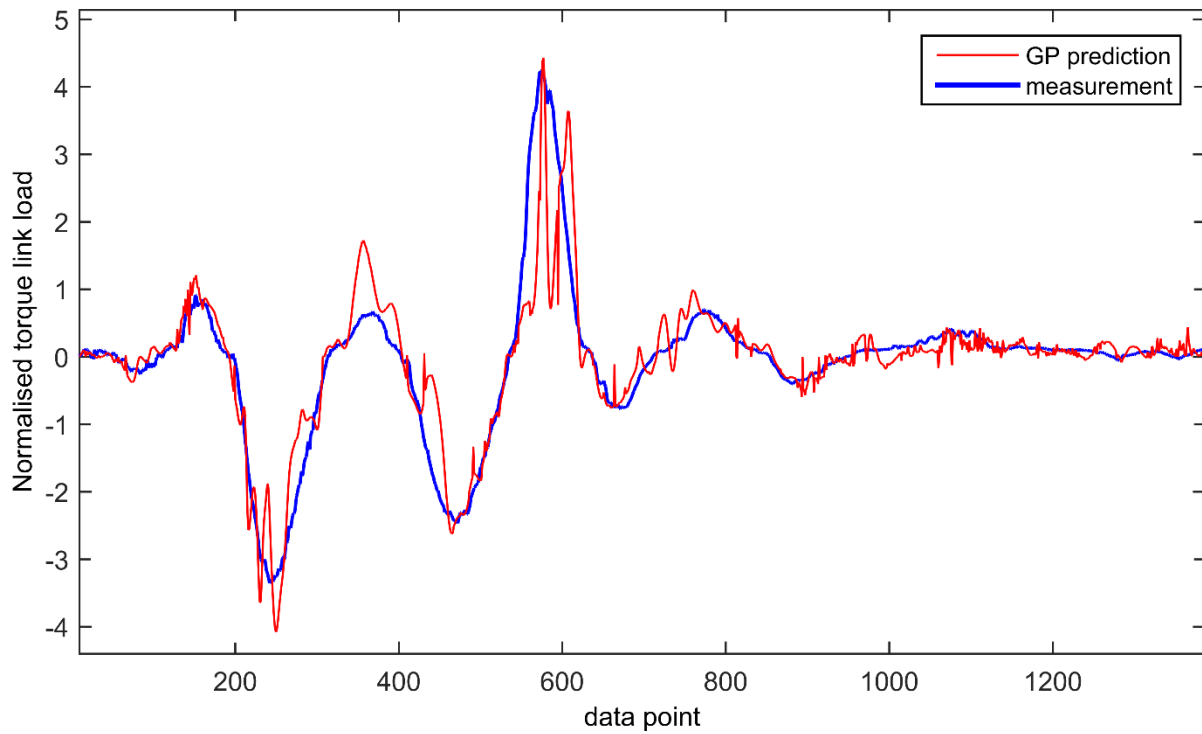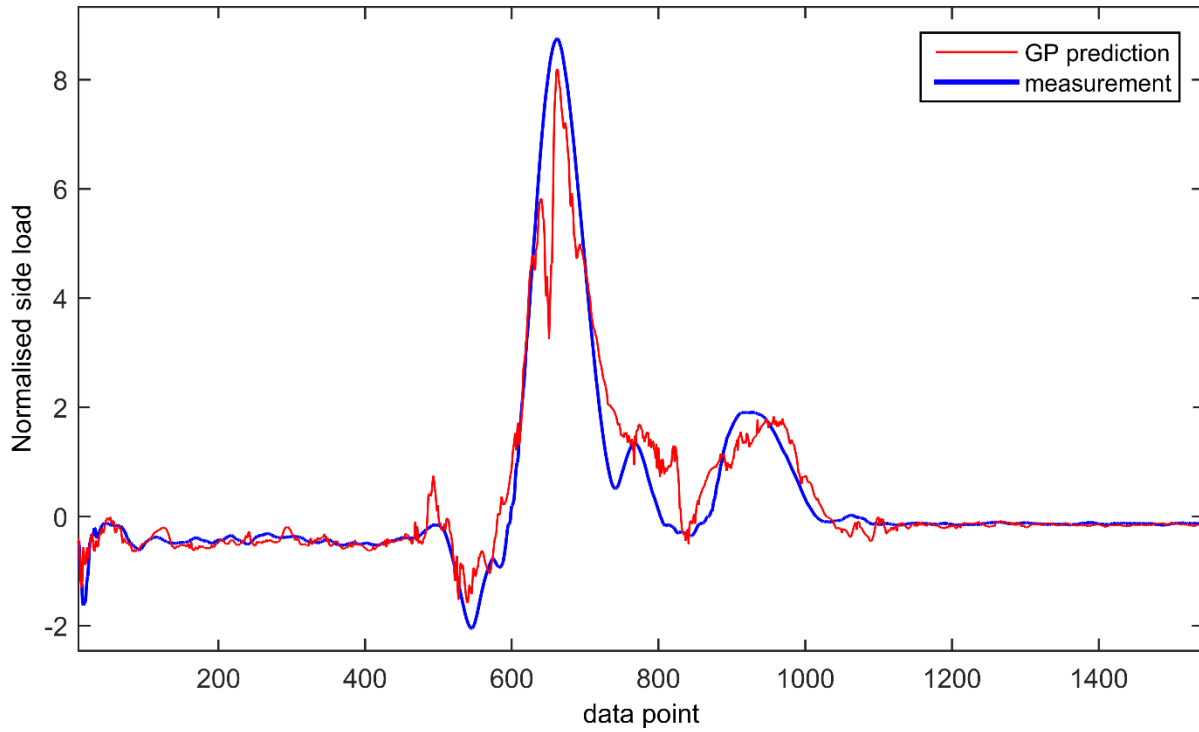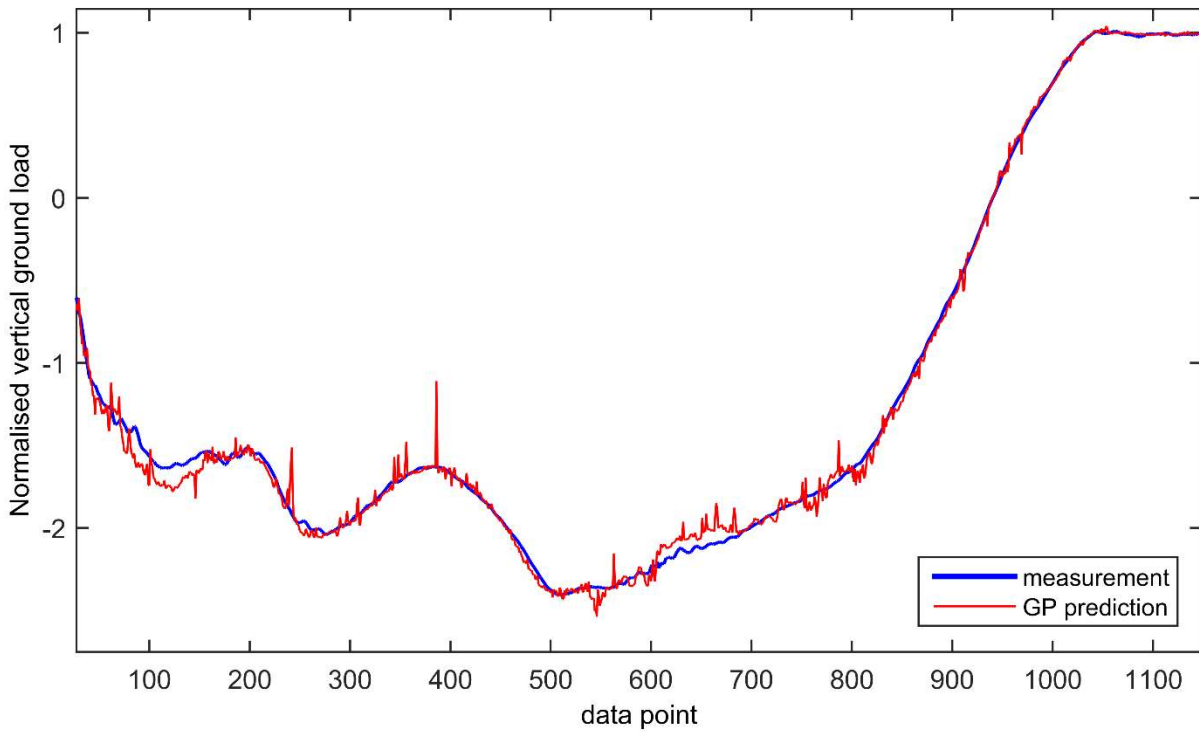
Figure 7 Prediction of torque-link load (unseen data) – inputs selected by greedy algorithm. Confidence intervals have been omitted for clarity.
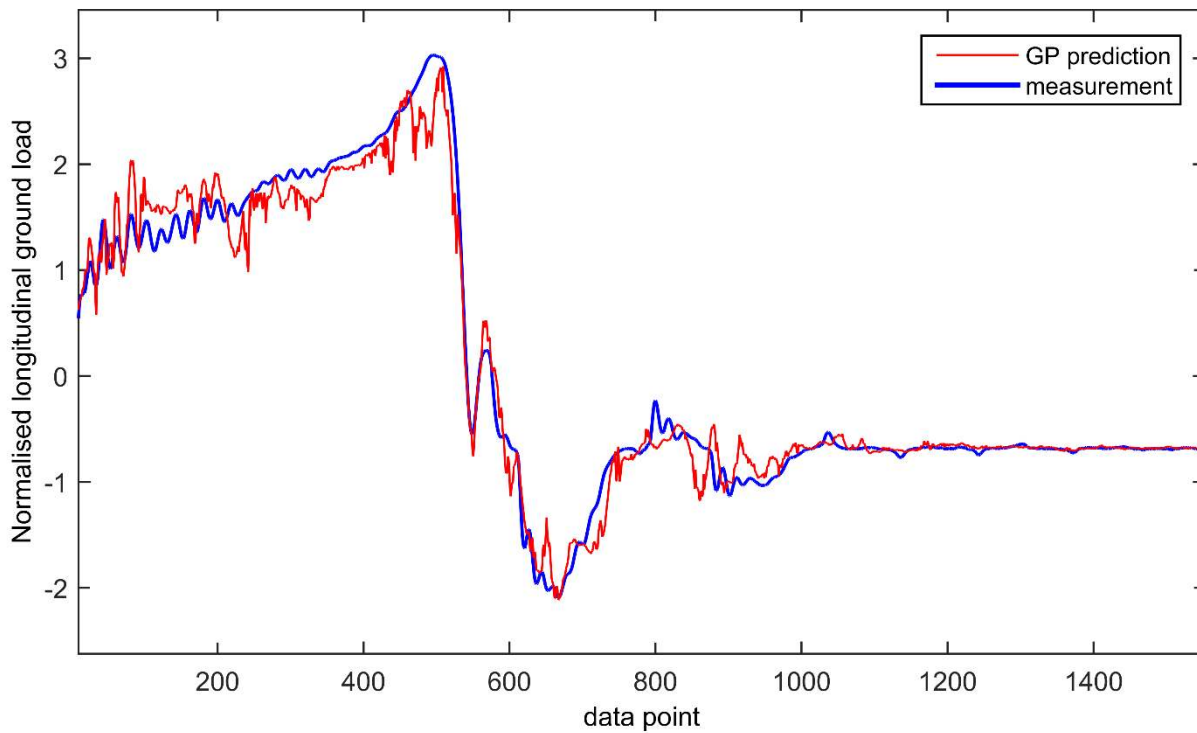
Figure 8. Prediction of drag ground-to-tyre load (unseen data) – inputs selected by greedy algorithm. Confidence intervals have been omitted for clarity

Figure 9. Prediction of side ground-to-tyre load (unseen data) – inputs selected by greedy algorithm. Confidence intervals have been omitted for clarity.

Figure 10. Prediction of vertical ground-to-tyre load (unseen data) – inputs selected by greedy algorithm. Confidence intervals have been omitted for clarity.

B. Flight Test Data

The previous section confirmed, in principle, the feasibility of making predictions for loads on landing gear from other measurements. In this section, the predictive capability of GP regression is explored in the more challenging context of measured flight test data. GPs are trained to predict vertical landing gear load given a number of input parameters recorded on the aircraft.

Prediction of Load Using Available Aircraft Measurements

A pertinent question here is that of whether landing gear loads experienced in operation can be predicted from measurements readily available on the aircraft today, such as data from the aircraft Flight Data Recorder. As an initial study, the choice of model inputs was hence limited to only those measurements more commonly available on aircraft, such as pitch, roll, yaw, and their rates, wheel speeds, brake pressures, altitude and acceleration at the cabin centre of gravity. At this stage a greedy algorithm was not employed for input variable selection.

For this trial, data from a fifth of the available landing events were used for model training and as for the drop test data this was subsampled using every fourth data point. The twenty-two measured variables were employed as model inputs. In this trial the results were not filtered. When the trained GP was tested on data from landing events not in the training set, the results were mixed. Across the testing set an average MSE of 47.5% was achieved, with the highest and lowest MSEs for single landings being 72.7% and 26.4% respectively. It was found that the predictions were generally able to follow the overall trend of the measured load. Figure 11 provides a typical comparison of a (normalised) measured load and the GP prediction, over a period of data in the testing set. The figure illustrates how the general trend is followed, along with where the model fails to predict the load accurately. An important aspect of employing Bayesian techniques such as GPs is also illustrated in this figure; towards the end of the time history, one can see that the confidence intervals expand away from the prediction. This indicates that the confidence in the prediction is lower, which is most likely due to the experiencing conditions in the input parameter set not present in the training set.
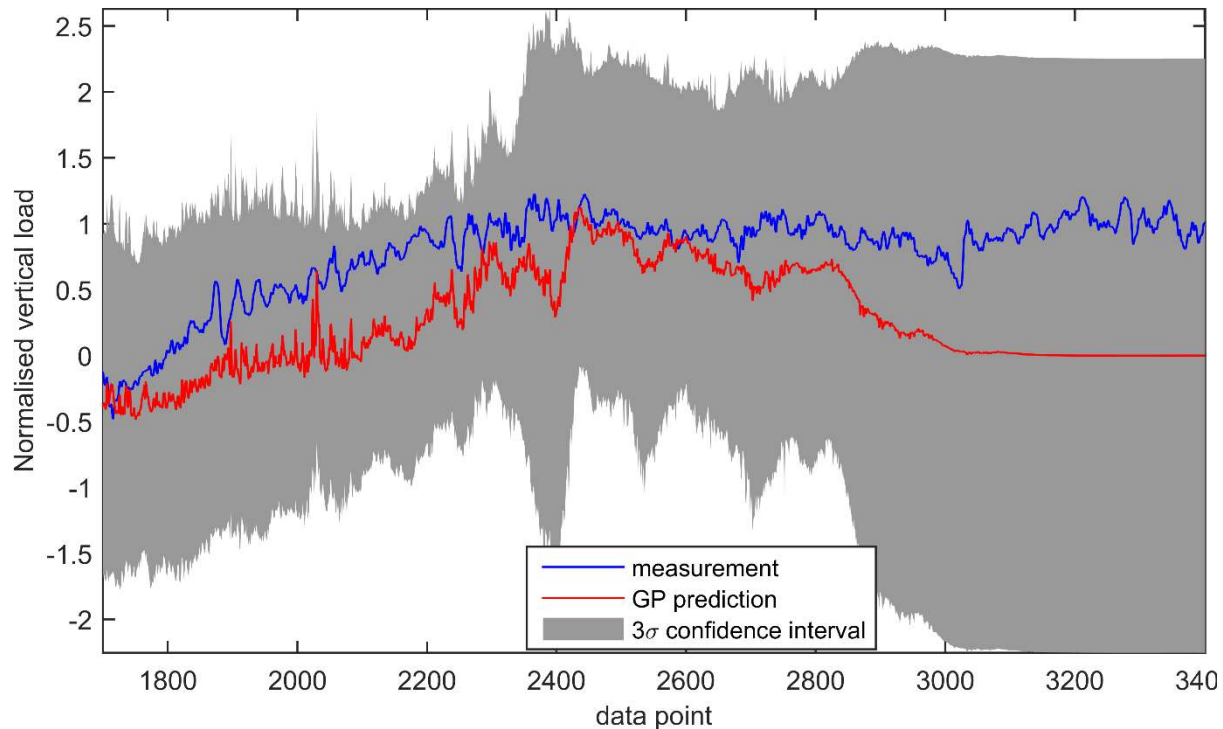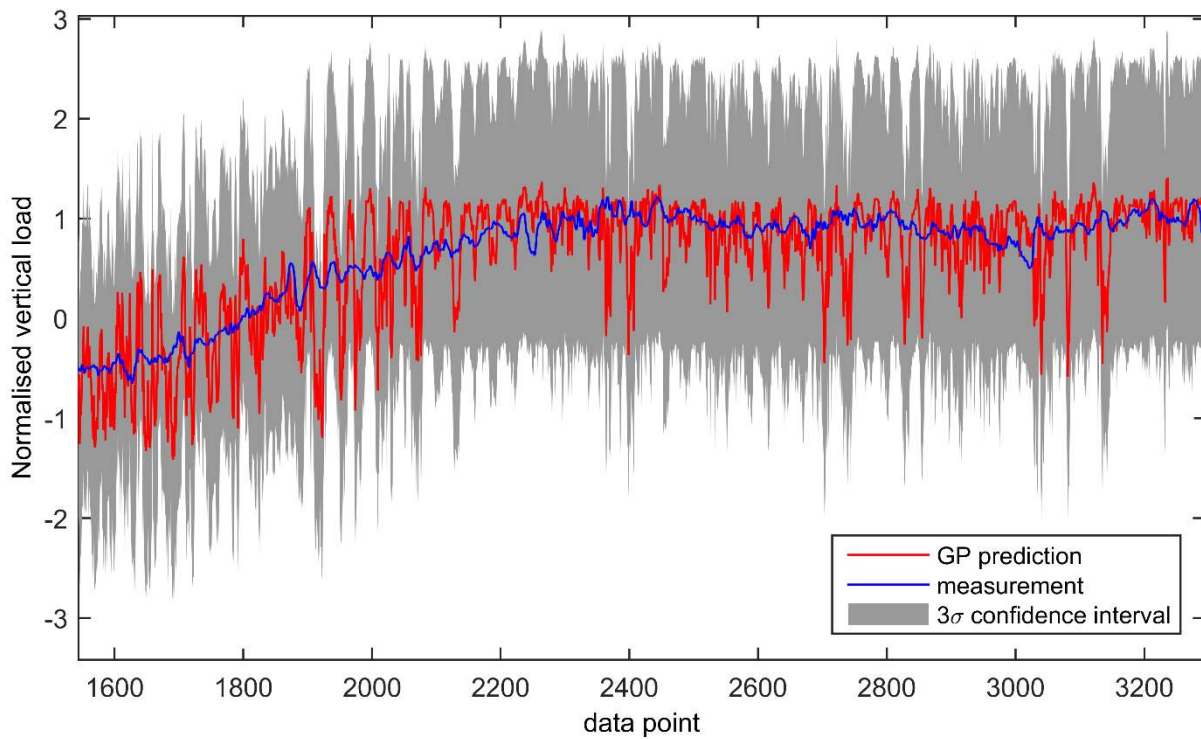
Figure 11. GP prediction capability for a sample test flight when using commonly available flight measurements.

When employing a greedy algorithm in an attempt to improve the model fidelity, interestingly, the only variables selected by the algorithm as suitable model inputs were the longitudinal, lateral and vertical accelerations measured at the cabin centre of gravity. Once these had been included as model inputs, no other variable would provide a decrease in MSE.

When employing only these acceleration measurements at the cabin centre of gravity as model inputs an average MSE of 38.4% was achieved across the testing set, with the highest and lowest MSEs for single landings being 73.7% and 17.5% respectively. With a reduced input set, the predictions generally provided a lower MSE. Figure 12 shows the GP prediction for the same test data as shown in Figure 11, when only measurement of the acceleration at the cabin centre of gravity is used as a model input. From this figure one can see that again the general trend of the load is followed by the prediction, but one can immediately see that

the predictions appear noisy. In some of the poorest predictions, this noise is a prohibitive issue, as can be seen in Figure 13, which shows the poorest prediction across the testing set. This is currently thought to be attributable to disparity between the sampling rates of the sensors measuring the load on the gear and the acceleration in the cabin.

A conclusion of this initial study is that useful information for prediction of loads on the landing gear is available from commonly measured flight parameters. However, the predictions in this study are variable with some being much more accurate than others. Further work is needed to ascertain whether prediction capability can be improved and by how much when only commonly measured input variables are used. In the next section of this work, model improvement is studied when input variables less commonly measured are included.

Figure 12. GP prediction capability when only using acceleration measurements at the cabin centre of gravity.
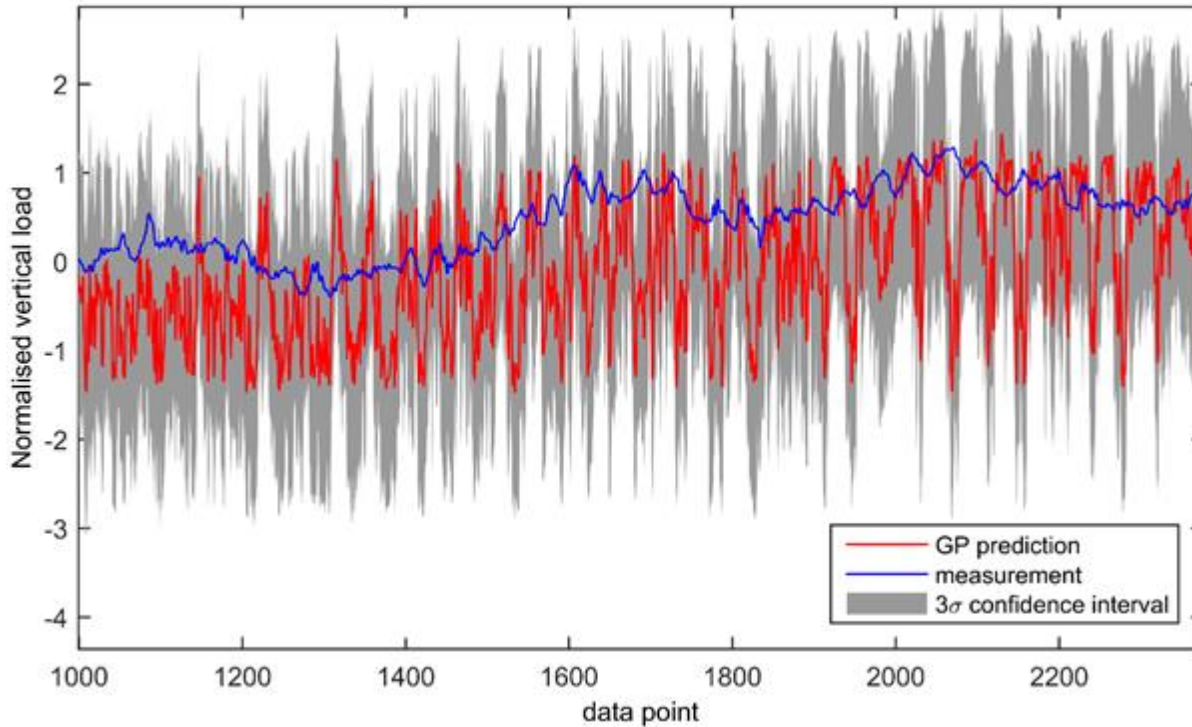
Figure 13. GP prediction capability when only using acceleration measurements at the cabin centre of gravity (poorest prediction across recorded landings).

**Prediction Load Using Additional Measurements**

Following on from results gained when working with drop test data, it was expected that additional predictive capability could be gained by using input variables that are not so commonly measured. Therefore, a number of additional measurements were included, such as acceleration measurements on the landing gear itself, shock absorber closures and gas pressures.

Initially, a GP was trained using this full set of inputs and with the same choice and subsampling of training data as in the last section. The mean MSE for the testing set was 6.3% with a range between 0.3% and 37.3%. This already shows a marked improvement on the results obtained using only the commonly measured quantities as inputs. Next, with the augmented set of candidate input parameters, a greedy algorithm was employed to select the model inputs and also the most informative landings for the training set. This latter addition was made in order to identify a training set that would provide a model that could generalise to the largest number of different landing conditions.

The termination criteria for the greedy algorithm applied was based on the reduction in MSE between each iteration. It was found here that after five iterations the overall test MSE was decreasing by a very small amount. The five input parameters selected by the greedy algorithm included the previously mentioned accelerations at the cabin centre of gravity, along with an acceleration measurement on the gear itself and measurements of shock absorber closure and gas pressure.

With these five inputs, the prediction capabilities on the test set were significantly improved. An average MSE of 3.5% was achieved across the testing set, with the highest and lowest MSEs for single landings being 6.3% and 1.4% respectively. An example of a typical GP prediction is shown in Figure 14, where one can easily see the increased predictive fidelity compared to Figures 11-13.

These results are encouraging, however, it must be noted that prediction capability is reduced around the initial touchdown period during landing. It is currently thought that this originates from the wider variability in conditions that affect this period in the landing, for example, the pilot control at this instant may vary
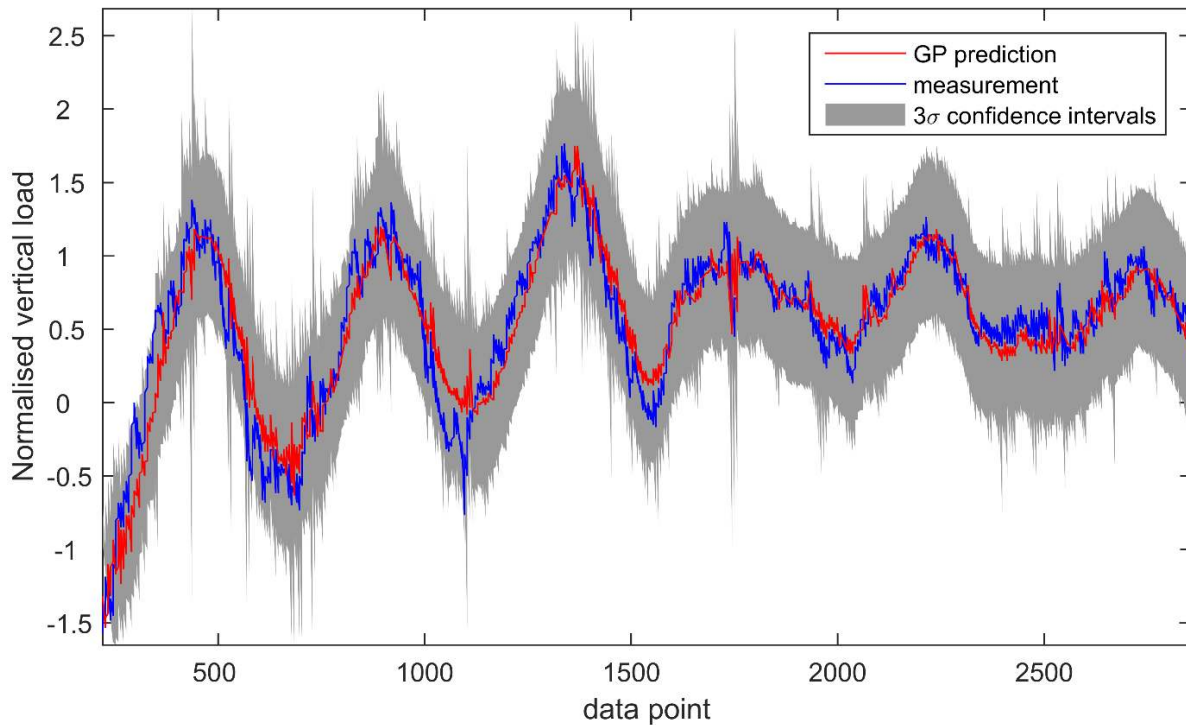
Figure 14. Improved GP prediction capability when employing additional input parameters not currently commonly measured.

significantly between different pilots and depending on weather conditions. Interestingly, these very accurate predictions were obtained without the need for filtering, as was necessary with the drop test data.

V.   CONCLUSION

GP regression has been used to create models for load prediction across the aircraft landing gear. Some of these models have been very successful at predicting the measured loads. GP regression has proved to be a very useful tool providing accurate predictions, and importantly confidence intervals on those predictions. Such confidence intervals have an inbuilt capability to signal if novel data unlike data previously seen is present, which will prove invaluable when predicting loads from any new data collected. Concerning flight test data, it has been found that when measurements from instrumentation on the gear itself have been included as model inputs, predictions of much higher accuracy are achievable on data from landings not in the training set. Where the input parameters employed were limited to only those measurements more commonly available (such as those dictated by Flight Data Recorder requirements), the prediction capability is reduced, however, in general model predictions are able to track the trend of recorded load in the gear.

The ideal possibility of prediction from commonly available measurements cannot yet be ruled out: a wider training data set and enhancements to the models and optimisation routines may yet yield sufficient modelling fidelity. However, it is possible that some extra instrumentation may be required in which case the subset selection methods trialled here give a way of searching for a minimal set of extra measurements. The industry requirements in terms of precision will also need to be taken into account and a possibility for future work is an optimisation routine which includes a trade-off between model precision and the cost of providing extra sensors.

The flight test data was all gathered from a single aircraft with a given landing gear configuration and differing runway surfaces. Further work will investigate further the variability arising from different conditions. It is likely, however, that any aircraft will require the training of a model that is particular to that aircraft. On the other hand, landing surface conditions, especially roughness, will have a direct effect

upon the inputs to the models and so a single model should be able to cope across different situations. Ongoing work is investigating how the noise parameter for the GP models can contain information about runway surface roughness.

Throughout this work, the question arose as to whether it was expedient to filter data. The question first arose during analysis of the drop test data, and in that case, filtering did indeed improve model predictions. The reason for the improvement was that the target outputs considered were much smoother than many of the input variables. This fact raised questions about the instrumentation; it appeared that the strain gauges used to measure loads may have had different frequency responses to the accelerometers used to measure accelerations. This is an important consideration; however, detailed information about the test instrumentation proved unobtainable. Interestingly, when working with flight test data, filtering was considered unnecessary, with prediction fidelity very good without it. In general, information about sensors and test signal conditioning (anti-aliasing filters etc.) will be important.

Selecting model inputs using a greedy algorithm proved to be a useful approach. Although greedy algorithms are well known to be suboptimal, they have been used here to avoid excessive computational cost incurred when using approaches that are known to be optimal, such as evolutionary algorithms. In the drop test data, where the number of inputs were limited, results when the using the greedy algorithm were not always better than when employing the whole range of available input parameters. However, neither were they significantly worse and were improved for some of the test conditions. This confirmed the feasibility of the approach. For the flight test data, model input selection is much more important because a wide range of measurements are possible. It is very encouraging, therefore, that the greedy algorithm results were a significant improvement on the models using the full candidate input set. True optimisation routines, such as a genetic algorithm, would potentially give further improvements if future computational resources allow.

By employing such models for predicting loads on the gear, the usage and ultimately the remaining life of the landing gear component can be predicted. However, in order to develop a landing gear parametric model, adequate 'training data' (i.e. historical data which can be used as a basis for 'learning'

the relationship between the flight parameters and the landing gear loads data) is required in all operating conditions. Therefore an effective operational loads monitoring capability is first required.

ACKNOWLEDGEMENTS

REFERENCES

1. Schmidt, R. K. and Sartor, P. (2009). "Landing Gear." *Encyclopaedia of Structural Health Monitoring.* Edited by C. Boller, F.-K. Chang, and Y. Fujino. Wiley. 1983-1994.
2. Chun-Yung Niu, M. (1999). *Airframe Structural Design.* Hong Kong Conmilit Press limited
3. Young, D.W. (1986). "Aircraft Landing Gear – Past, Present and Future." *Proceedings of the Institution of Mechanical Engineers Part D Transport Engineering.* 200 (24): 75-92
4. Schmidt, R. K. (2009). "System and Method for Determining Aircraft Hard Landing Events from Inertial and Aircraft Reference Frame Data." *International Patent WO 2006/130984.*
5. Yamada, M. and Kawakami, Y. (2008). "Landing Gear MEMS Health Monitoring Platform, Sumitomo Precision Products Co., Ltd." *SAE A-5 Aerospace Landing Gear Systems Conference.* Toulouse, France.
6. Woodward, S. E., Coffey, N. C., Gonzalez, G. A., Taylor, B. D., Brett, R. R., Woodman, K. L., Weathered, B. W., and Rollins, C. H. (2004). "Development and Flight Testing of an Adaptable Vehicle Health Monitoring Architecture." *Journal of Aircraft.* 41 (3): 531-539.
7. Cowan, S. J., Cox, R. L., Slusher, H. W., and Jinadasa, S. (2004). "Airplane Hard Landing Indication System." *United States Patent 6 676 075.*
8. Schmidt, R. K. (2008). "Monitoring of Aircraft Landing Gear Structure." *The Aeronautical Journal.* 112(1131): 275-278.
9. Dellac, S. and Lafaye, E. (2006). "Force-Measurement Cell and a Connection Pin Fitted with such a Cell." *United States Patent 2006/0266561.*
10. Kehlenbeck, U. (1999). "Airbus A340 Weight and Balance System." *58th Annual Conference of Society of Allied Weight Engineers, Society of Allied Weight Engineers, Inc.* San Jose, California. 2450.
11. Nelson, H. K., Kleingartner, C. A., and Vetsch, L. E. (1981). "Strain/Deflection Sensitive Variable Reluctance Transducer Assembly." *United States Patent 4 269 070.*
12. Patzig, H.-N. and Schult, K. (1993). "Arrangement of Sensors on the Landing Gear of an Aircraft for Measuring the Weight and Position of Center of Gravity of the Aircraft." *United States Patent 5 257 756.*
13. Nance, K. (2001). "Method of Determining Status of Aircraft Landing Gear." *United States Patent 6 293 141.*
14. Yates, M. S. and Keen, P. (2007). "Landing Load Monitor for Aircraft Landing Gear." *International Patent WO 2007/023280.*
15. Azzam, H. (1997). "A practical approach for the indirect prediction of structural fatigue from measured flight parameters." *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 211(G1): 29-38.
16. Azzam, H., Hebden, I., Gill, L., Beaven, F. and Wallace, M. (2005). "Fusion and Decision Making Techniques for Structural Prognostic Health Management." *IEEE Aerospace Conference Paper #1535.*
17. Wallace, M., Azzam, H. and Newman, S. (2004). "Indirect approaches to individual aircraft structural monitoring." *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering.* 218: 329-346.
18. Reed, S. (2004). "The Use of Artificial Neural Networks in Aircraft Fatigue Monitoring - Prediction of Undercarriage Strains." Tech. Rep. QinetiQ/FST/CR041003. QinetiQ.
19. Reed, S. (2006). "Indirect Aircraft Structural Monitoring Using Artificial Neural Networks." *Ph.D. thesis.* Department of Mechanical Engineering, University of Sheffield.
20. Mohanty S., Chattopadhyay, A., Peralta, P. and Das, S. (2011). "Bayesian Statistic Based Multivariate Gaussian Process Approach for Offline/Online Fatigue Crack Growth Prediction." *Experimental Mechanics.* 51(6): 833-843E.
21. E. J. Cross, K.Y. Koo, J.M.W. Brownjohn and K. Worden. (2012). "Filtering environmental loading effects to enhance novelty detection on cable-supported bridge performance." *6th International conference on Bridge Maintenance, Safety and Management.* Lake Maggiore, Italy.
22. Hensman, J., Mills, R., Pierce, S.G., Worden, K., and Eaton, M. (2010). "Locating acoustic emission sources in complex structures using Gaussian processes." *Mechanical Systems and Signal Processing.* 24: 211-223.

23. Hensman, J., Pullin, R., Eaton, M., Worden, K, Holford, K.M., and Evans, S.L. (2009). "Detecting and identifying artificial acoustic emission signals in an industrial fatigue environment." *Measurement Science and Technology*. 20, 045101.
24. Cross EJ, Sartor P, Worden K & Southern P (2012) Prediction of Landing Gear Loads Using Machine Learning Techniques. *Structural Health Monitoring* 2012 (pp 1056-1063)
25. Cross EJ, Sartor P, Worden K & Southern P (2013) Prediction of Landing Gear Loads from Flight Test Data Using Gaussian Process Regression. *Structural Health Monitoring* 2013, vols 1 and 2 (pp 1452-1459)
26. Messier-Bugatti-Dowty Design Guide. (2011). "Drop, Free Extension and Gas Spring Verification Tests and Correlation." DSG-4208.
27. Messier-Bugatti-Dowty Qualification Test Report, QTR00015-1. (2009).
28. Wright, J. R. and Cooper, J. E. (2007). *Introduction to Aircraft Aeroelasticity and Loads*. Wiley.
29. Currey, N. S. (1988). *Aircraft Landing Gear Design: Principles and Practices*. American Institute of Aeronautics and Astronautics.
30. Rasmussen, C.E., and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning*. MIT Press
31. Worden, K. and Burrows, A.P. (2001). "Optimal sensor placement for fault diagnosis." *Engineering Structures*. 23: 885-901.
32. Akaike, H. (1974) "A new look at the statistical model identification." *IEEE Transactions on Automatic Control*. 19: 716-723.
33. Naish-Guzman, A. and Holden, S. (2007). "The Generalized FITC Approximation." *Advances in Neural Information Processing Systems*. 20.